Medical Science

25(110), April, 2021

The distribution of Duffy alleles and genotypes in a Western Saudi cohort

Waleed M Bawazir[1,2]✉, **Asim Khogeer**[3], **Azoof Ashour**[1], **Mohammad Sabbag**[4], **Raed Alserihi**[1,6], **Raed Felimban**[1,6], **Nora Hakami**[1], **Malik Altayar**[5], **Ahemd Mirza**[1], **Refaat Elfayomi**[1], **Thoraia Shinawi**[1]

To Cite:
Bawazir WM, Khogeer A, Ashour A, Sabbag M, Alserihi R, Felimban R, Hakami N, Altayar M, Mirza A, Elfayomi R, Shinawi T. The distribution of Duffy alleles and genotypes in a Western Saudi cohort. *Medical Science*, 2021, 25(110), 999-1009

Author Affiliation:
[1]Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia
[2]Hematology Research Unit, King Fahd Medical Research Center, King Abdulaziz university, Jeddah, Saudi Arabia
[3]Research and Plan Department, General Directorate of Health Affairs Makkah Region, Saudi Arabia Ministry of Health, Makkah, Saudi Arabia
[4]Regional Laboratory and Blood Bank, Health Affairs Makkah Region, Saudi Arabia Ministry of Health, Makkah, Saudi Arabia
[5]Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, University of Tabuk, Tabuk, Saudi Arabia
[6]Center of Innovation in Personalized Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

✉**Corresponding author**
Department of Medical Laboratory Sciences, Faculty of Applied Medical Sciences, and Hematology Research Unit,
King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia.
Email: wbawazir@kau.edu.sa

Peer-Review History
Received: 12 March 2021
Reviewed & Revised: 13/March/2021 to 20/April/2021
Accepted: 21 April 2021
Published: April 2021

Peer-review Method
External peer-review was done through double-blind method.

ABSTRACT

Background and Objective: The Duffy blood group is clinically significant in transfusion medicine, and is based on the presence of glycoproteins known as Fy antigens on the surface of red blood cells. In Saudi Arabia, serological studies on blood groups, including Fy antigens have been undertaken, however the population distribution of *FY* alleles and genotypes requires further elucidation. Therefore, we analysed samples from random Western Saudis blood donors to establish the frequency of *FY* alleles and genotypes in this cohort. *Materials and Methods*: Touchdown PCR and DNA sequencing was used to analyse the following single nucleotide polymorphisms, c.125G>A, c.1-67T>C and c.256C>T, in 349 random blood donors from Jeddah and Makkah. The c.298G>A polymorphism was used to determine the allele designation of weak *FY* alleles. *Results*: Our investigation showed no significant differences between Makkah and Jeddah blood donors, except for the c.125G>A polymorphism, where the A allele was significantly higher in Jeddah donors (odds ratio (OR) = 1.56 (1.09–2.24)). Most donors (39%) had the homozygous *FY*02N.01/FY*02N.01* genotype, similar to African populations. Seven individuals had the *FY*01N.01/FY*02N.01* genotype. Five had the predicted genotypes of *FY*02/FY*02W.01, FY*01/FY*02W.01, FY*01/FY*01W.02* and *FY*02W.01/FY*02N.01*. The remaining donors had predicted genotypes that expressed Fy antigens on the erythrocyte surface. *Conclusion*: The molecular genetic background of Western Saudis is similar to African populations, however in terms of allele and genotype frequencies, Western Saudi Arabia contains a unique genetic pool, distinct to other populations.

Keywords: Blood groups, Donors, Genetics, Genotyping

1. INTRODUCTION

The Duffy blood group system was identified in a 43-year-old patient with haemophilia, who received several red blood cell (RBC) units, and subsequently developed jaundice after transfusion with anti-RBC antibodies (Cutbush & Mollison, 1950a,b). Duffy is an N-glycosylated multi-pass

transmembrane glycoprotein, composed of 336 amino acid residues. The Duffy (*FY)* gene is localised to chromosome 1, and encodes a glycoprotein (GP) that serves as a Duffy antigen receptor for chemokines (DARC), commonly called atypical chemokine receptor 1 (ACKR1 or CD234). As a chemokine receptor, it is attracted to cell sites undergoing inflammation, and enrols other blood cell types to the area. These molecules include, acute inflammation chemokines (CXRs), chronic inflammation chemokines (the CC class), interleukin 8 (IL-8), and regulated upon activation normal T-cell expressed and secreted (RANTES) (Mohandas, 2005). In animal models, mice lacking the *FY* gene were no more susceptible to infection than groups that expressed *FY,* suggesting the function of the Duffy chemokine receptor is unclear (Luo et al., 2003).

Currently, it is accepted that the two Duffy mRNA transcripts encode for a minor, less abundant alpha form, comprising a 338 amino acid Fy glycoprotein, and a 336 amino acid protein called the major beta form. Approximately, 10,000 Duffy protein copies are found on the surface of RBCs, and are involved in chemokine binding processes (Pogo & Chaudhuri, 2000). The Duffy protein is also found on epithelial cells in kidney collecting ducts, Purkinje cells of the cerebellum, lung alveoli and the endothelial cells of capillary and postcapillary venules (Rot & Von-Adrian, 2004). While the exact role of the Duffy glycoprotein is not precisely understood, it may play a role in inflammation and malaria infections, by acting as an erythroid receptor for *Plasmodium vivax* and *P. knowlesi* via Fy epitope binding (Chitnis et al., 1996). Therefore, a Duffy protein lacking a receptor binding site, or the absence of Duffy antigens on RBCs makes these cells more resistant to invasion from a malarial species (Hadley et al., 1986). The receptor-binding site for chemokines, the parasite-specific binding site, and the major antigenic domain sites are placed in overlapping regions at the extracellular N-terminal terminus.

The Duffy blood group system has six known antigens (Fy^a, Fy^b, Fy^3, Fy^4, Fy^5, Fy^6), with four major phenotypes; Fy(a+b+), Fy(a−b+), Fy(a+b−) and Fy(a−b−). The main two antigen types, Fy^a (FY1) and Fy^b (FY2) differ in a single amino acid at position 42 in the extracellular domain, resulting in Fy^a expression (glycine change) and Fy^b expression (aspartate change) (Chaudhuri et al., 1989; Pogo & Chaudhuri, 2000; Hughes et al., 2007). However, the rare Duffy antigen variant, Fy^3 appears to be of clinical significance and carried on DARC (Reid & Mohandas, 2004; Olteanu et al., 2005; Meny, 2010). Fy^a and Fy^b are relatively frequent in Europeans i.e. Fy^a occurs in 39.8% and Fy^b in 60.2%, in Africans, Fy^a occurs in 1.9% and Fy^b in 98.1%. In East-Asians, Fy^a occurs in 92.3% and Fy^b in 7.7% (Höher et al., 2018). Duffy antibodies against Fy^a and Fy^b antigens are connected as the cause of a transfusion reaction (Le Pennec et al., 1987; Talano et al., 2003; Kim et al., 2004).

At the molecular genetic level, *FY* consists of two exons distributed over 1500 bp of genomic DNA (Lee et al., 1974). A single nucleotide substitution at position 125 (c.125G>A) represents the distinguishing feature between the *FY*A* (*FY*01*) and *FY*B* (*FY*01*) alleles (Iwamoto et al., 1996). Two gene products generate the Duffy negative phenotype; Fy(a-b-) (Daniels, 2005). Most commonly known is the *c.1-67T>C* mutation in the promoter region of the *FY* allele, *FY*02N.01* suppresses Duffy glycoprotein expression in RBCs, however the protein still functions in other cell types (Rios et al., 2000). This mutation is found in approximately 70% of African American populations, and nearly 100% of West African individuals (British Committee for Standards in Haematology, 2013). The single nucleotide polymorphism (SNP) *c.256C>T*, with or without *c.298G>A*, is associated with the weak expression of Duffy glycoproteins (Olsson et al., 1998). A recent report showed that this SNP was not only associated with *FY*B,* but also with the *FY*A* allele (Lopez et al., 2015).

A blood bank centre delivers appropriate blood products to vulnerable individuals who require blood transfusions (Reesink et al., 2008). However, this is challenging if blood recipients have alloantibodies in the absence of important information of donor phenotypes (Owaidah et al., 2020). A recent study indicated that the majority of the Saudi Arabian population has the Fy (a-b-) phenotype (Lopez et al., 2015). However, most studies on blood group distributions in Saudi Arabia have been conducted using serological methods (Bashwari et al., 2001; Elsayid et al., 2015, Alabdulmonem et al., 2020). This discrepancy can be resolved by analysing blood samples from different parts of Saudi Arabia, using molecular techniques such as touchdown PCR and DNA sequencing, to identify Duffy genetic backgrounds, including the null phenotype, Fy(a-b). To date, *FY* genotyping and its associated polymorphisms in Saudi Arabian populations has not been conducted, therefore we investigated the genetic background, distribution of main *FY* alleles and genotypes, including weak Duffy genotypes, in a Western Saudi cohort.

2. MATERIALS AND METHODS

Patient samples

EDTA mixed blood samples were collected from 349 randomly selected, healthy donors from Jeddah (n = 220) and Makkah (n = 129). The male to female ratio was 1:0.93 (Table 1). Sample collection was performed at the blood bank of King Abdullah Medical Complex, King Fahd General Hospital, and King Fahd Armed Hospital (Jeddah, Saudi Arabia). Two blood banks at Alnoor Hospital and the Regional Laboratory (Makkah, Saudi Arabia) were used as sources of the Meccan blood donors. Male blood

donors were aged between 20 – 60 years old, and female blood donors were aged between 20–70 years old. All participants signed consent forms. Samples were processed, separated and stored in King Abdul-Aziz University and King Fahd Research Center. All samples were collected between November 2016 and January 2019. No inclusion or exclusion criteria were implemented.

Table 1 Participating blood donor numbers

City	No. of samples	Male	Female	M:F ratio
Jeddah	220	113	107	1–0.95
Makkah	129	68	61	1–0.9
Total	349	181	168	1–0.93

DNA extraction

DNA was extracted from whole blood or buffy coat samples using a commercial blood extraction genomic DNA isolation Kit (Norgen Biotek Corp, Ontario, Canada). Briefly, cells were lysed by adding 150 µl whole blood, an equal volume of digestion buffer, and 12 µl proteinase K to a tube, followed by incubation at 55 °C for 60 min. Then, samples were vortexed for 20 s before 200 µl buffer SK was added. Then, 300 µl ethanol was added and the mixture transferred to a spin column inside a collection tube. The tube was centrifuged at 5200 g for 3 min, after which the column was washed twice in 500 µl wash solution A, before moving the column into a new tube. DNA elution occurred by adding 200 µl elution buffer B, and centrifuging the tube at 14000 g for 1 min. The elution volume was collected after a second elution at 14000 g for 2 min. DNA was quantified using a Nano Drop spectrophotometer (Nano Drop Technologies, Wilmington, DE, USA) at 260/280 nm. DNA concentrations were calculated using the mean of two measurements.

Duffy blood group genotyping

Duffy gene exons were amplified using PCR primers (Table 2). Touchdown PCR was conducted in a 25 µl PCR volume using GoTaq® Green Master Mix (Promega, Madison, Wisconsin, USA). An annealing temperature of 52°C was used over 35 cycles. PCR products were uni-directionally sequenced by cycle sequencing on an ABI 3730x (Applied Biosystems, Carlsbad, CA, USA). Alternatively, samples for DNA sequencing were outsourced to Macrogen (Seoul, South Korea). Raw DNA sequencing data were analysed using Geneious software (Biomatters Inc., Newark, NJ, USA).

Table 2 PCR primer sequences

Amplification Region	Primer Name	Forward Primer	Reverse Primer	PCR Product Size (Base pairs)
Promoter and Exon 1 regions	Duffy Pro	5′-GAG TGT AGT CCC AAC CAG CC-3′	5′-GGG AAA TGA GGG GCA TAG GG-3′	300
Exon 2	Fy2-1	5′-TAG CTC CCT GTG TCC CTC-3′	5′-ATG GCA CCC TAG CAG CAA-3′	300
	Fy2-2	5′-CTT TGC TAG GGT GCC AT-3′	5′-AAG AAA CCA CCC TCA CAA A-3′	300
	Fy2-3	5′CAA CCA GCC AAA TCC AAC CT-3′	5′-AAG AGG GAG CTA GGA GGC TA-3′	300

Statistical analysis

The frequency of Duffy alleles were analysed using an odds ratio method to determine whether the allele distribution in Jeddah blood donors were similar to the blood donors of Makkah. In addition, two-tailed Chi-Squared or Fisher-Exact tests were also used for the same objective. Average duffy allele and genotype frequencies of all population from the 1000 genomes project (https://www.internationalgenome.org) was used to examine whether it is similar to the duffy alleles and genotypes in this study. For this purpose, two-tailed Chi-squared or Fisher exact tests were used. Statistical analyses were conducted using GraphPad Prism software (La Jolla, California, USA). A $P - value$ of 0.05 was considered statistically significant.

3. RESULTS

Allele and genotype frequencies in samples

The allele and genotype frequencies of 349 samples are shown (Table 3, Figure 1a - b). The overall allele frequency for the *c.1-67T>C* polymorphism indicated that the T allele was slightly higher than the C allele in samples (57.4% and 42.6%, respectively) (Figure 1a). On the other hand, the A allele of the *c.125G>A* polymorphism was approximately 3.5 times higher than the G allele (A = 77.5%; G = 22.5%) (Figure 1a). The frequency of the C allele in the *c.256C>T* polymorphism was approximately 99.3% in samples, whereas the T allele frequency was 0.007% (Figure 1a).

The overall genotype distribution of the *c.1-67T>C*, *c.125G>A*, and *c.265C>T* polymorphisms in Western Saudis is shown (Table 3, Figure 1b). Analysis of the *c.1-67T>C* SNP showed that 40.7% of blood donors expressed the CC genotype, which abolished Fy antigen expression (*FY*Null*). The remaining samples expressed either heterozygous or wild type genotypes for the *c.1-67T>C* polymorphism (TC = 33.5% and TT = 25.8%) (Table 3, Figure 1b). As for the *c.125G>A* polymorphism, we observed a dominance of the A/A genotype (*FY*B*) in 63.9% of samples, whereas the heterozygous GA genotype (*FY*AB*) was found in approximately one third of the study population (27.2%), and the homozygous GG genotype (*FY*A*; 8.9%) (Table 3, Figure 1b). Only five samples (1.4%) were heterozygous for the *c.265C>T* polymorphism, whereas 98.6% of samples had the wild type genotype, CC. No samples were homozygous for the mutant form of the *c.256C>T* polymorphism (Table 3, Figure 1b).

Table 3 Allele and genotype frequencies of Duffy blood groups in blood donors from Western Saudi Arabia

Polymorphism	Allele	Genotype	Allele/genotype frequency			OR (95% CI)	*P - Value**
			All blood donors n (%)	Jeddah blood donors n (%)	Makkah blood donors n (%)		
c.1-67T>C	T		401 (57.4)	257 (58.4)	144 (55.8)	0.87 (0.65–1.19)	0.53
	C		297 (42.6)	183 (41.6)	114 (44.1)		
		TT	90 (25.8)	54 (24.5)	36 (27.9)		
		TC	117 (33.5)	75 (34.1)	42 (32.6)		0.79
		CC	142 (40.7)	91 (41.4)	51 (39.5)		
c.125G>A	G		157 (22.5)	86 (19.5)	71 (27.5)	1.56 (1.09–2.23)	0.019
	A		541 (77.5)	354 (80.5)	187 (72.5)		
		GG	31 (8.9)	15 (6.8)	16 (12.4)		0.09
		GA	95 (27.2)	56 (25.5)	39 (30.2)		
		AA	223 (63.9)	149 (67.7)	74 (57.4)		
c.265C>T	C		693 (99.3)	436 (99.1)	257 (99.6)	0.57 (0.09–3.61)	0.66
	T		5 (0.7)	4 (0.9)	1 (0.4)		
		CC	344 (98.6)	216 (97.7)	128 (99.2)		0.66
		CT	5 (1.4)	4 (1.8)	1 (0.8)		
		TT	0 (0)	0 (0)	0 (0)		

*Either Pearson's chi squared or Fisher Exact test P-value was used to examine differences between allele distributions in the two blood donor groups.

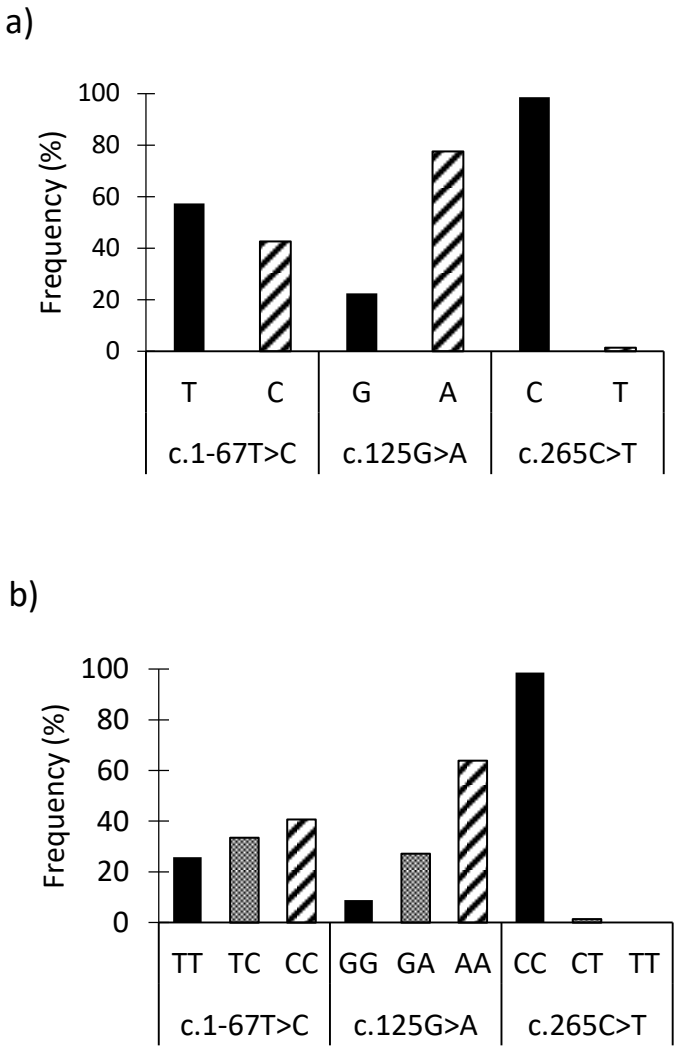

Figure 1 demonstrates the overall *FY* alleles and genotypes frequencies found in the Western blood donors of Saudi Arabia. (a) The overall allele frequency for the c.1-67T>C polymorphism indicated that the T allele was slightly higher than the C allele in samples (57.4% and 42.6%, respectively). As for the c.125G>A polymorphism, the A allele of the was approximately 3.5 times higher than the G allele (A = 77.5%; G = 22.5%). Expectedly, the majority of blood donors had the c.256C allele (99.3%). (b) The overall genotype distribution of the c.1-67T>C, c.125G>A, and c.265C>T polymorphisms demonstrated that the majority of blood donors were homozygous for the c.1-67T>C genotype (40.7%), which abolished *FY* antigen expression (*FY*Null*). The remaining blood donors had either heterozygous or wild type genotypes (TC = 33.5% and TT = 25.8% respecitively). As for the c.125G>A, a dominance of the AA genotype in 63.9% of blood donors, whereas the heterozygous GA and GG genotype were found the remaining blood donors (27.2% and 8.9% respectively). When it comes to the c.265T>C polymorphism, the majority of blood donors were homozygous for the CC genotype (98.6%) and few blood donors had the heterozygous genotypes TC (1.4%) but none of the blood donors had a homozygous CC genotype.

We also analysed differences in Duffy polymorphism allele and genotype frequencies between blood donors from Jeddah (J) and Makkah (M) (Table 3, Figure 2a - b). Analysis of the *c.1-67T>C* polymorphism in both groups revealed a high frequency of the T allele (J = 58.4% and M = 55.8%, respectively), in comparison to the C allele (J = 41.6% and M = 44.1%). However, this difference was not statistically significant (*P - value* = 0.53, OR = 0.87 [0.65 – 1.19]) (Table 3, Figure 2a), suggesting that C and T allele frequencies were similar in both groups. Similarly, we observed no significant differences in the frequency of *c.1-67T>C* genotypes in both groups (*P - value* = 0.79) (Table 3, Figure 2b). Analysis of the *c.125G>A* genotype demonstrated a dominance of the A allele, and AA genotype (OR = 1.56 [1.09–2.23]) (Table 3, Figure 2a - b). Interestingly, the A allele was significantly higher in Jeddah samples when compared to Makkah samples (*P - value* = 0.019). The AA genotype was dominant in both groups (67.7% and 57.4%, respectively) (Table 3, Figure 2b). For the c.265C>T polymorphism, the C allele and the CC genotype were dominant in both groups, but with no significant differences (*P- value* = 0.66). Similarly, neither group demonstrated a homozygous TT genotype (Table 3, Figure 2a – b & figure 3).

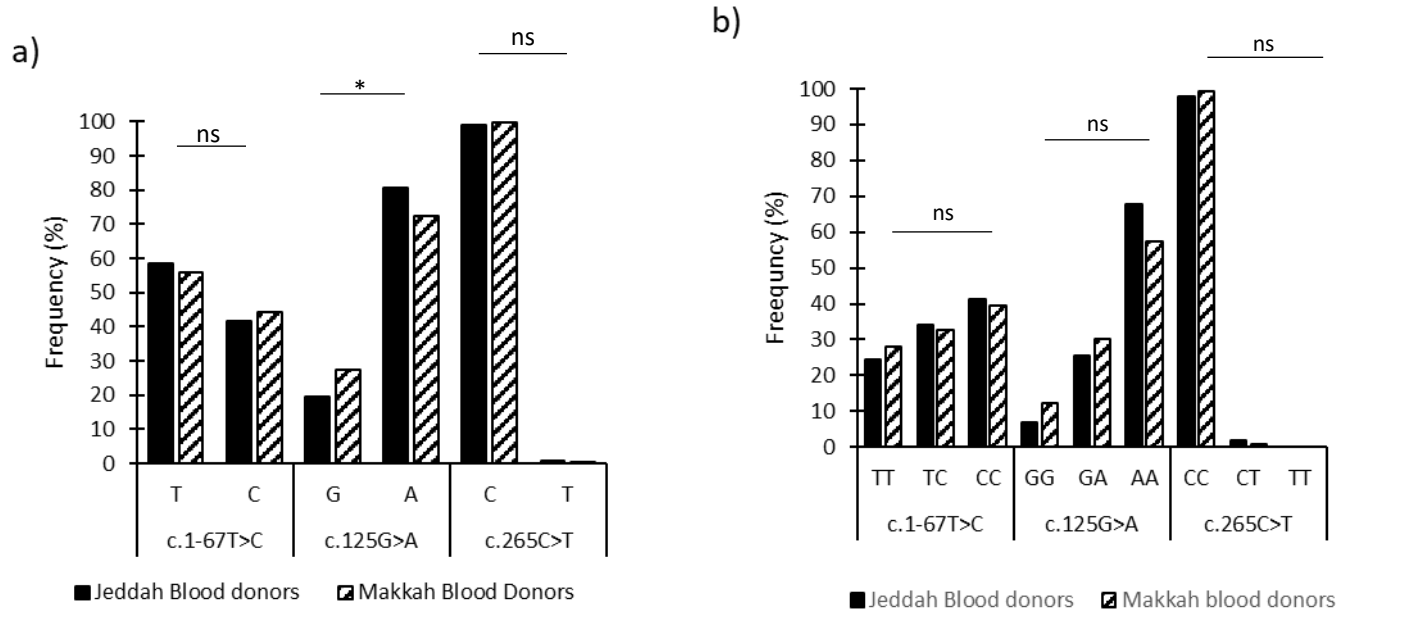

Figure 2 shows the comparisons between the *FY* alleles and genotypes frequencies found in Jeddah and Makkah blood donors of Western Saudi Arabia. (a) A comparison of the allele frequencies between Jeddah and Makkah blood donors for the c.1-67T>C, c.125G>A, and the c.265C>T single nucleotide polymorphisms were not statistically significant except for the c.125G>A (Odds ratio = 1.56 [95% CI = 1.09–2.23]; *P – value* = 0.019). (b) The genotype frequencies in the Jeddah blood donors were not significantly different from the genotype frequncies reported in Makkah blood donors (*P – value > 0.05*). Odds ratio was used to examine whether there was no difference in *FY* allele frequencies between Jeddah and Makkah blood donors. Analyses of the contingency tables of *FY* alleles and genotypes frequencies found in Jeddah and Makkah blood donors were conducted using two-tailed 2X2 or 3X2 chi-squared test to examine whether there was no difference in the allele and genotype frequencies in both groups of blood donors. Alternatively, two-tailed 2X2 fisher exact test was used when the contingency table contains an allele or genotype frequency of less than 5. A *P – value* of less than 0.05 was considered statistically significant. (*) statistically significant difference; (ns) not statistically significant.

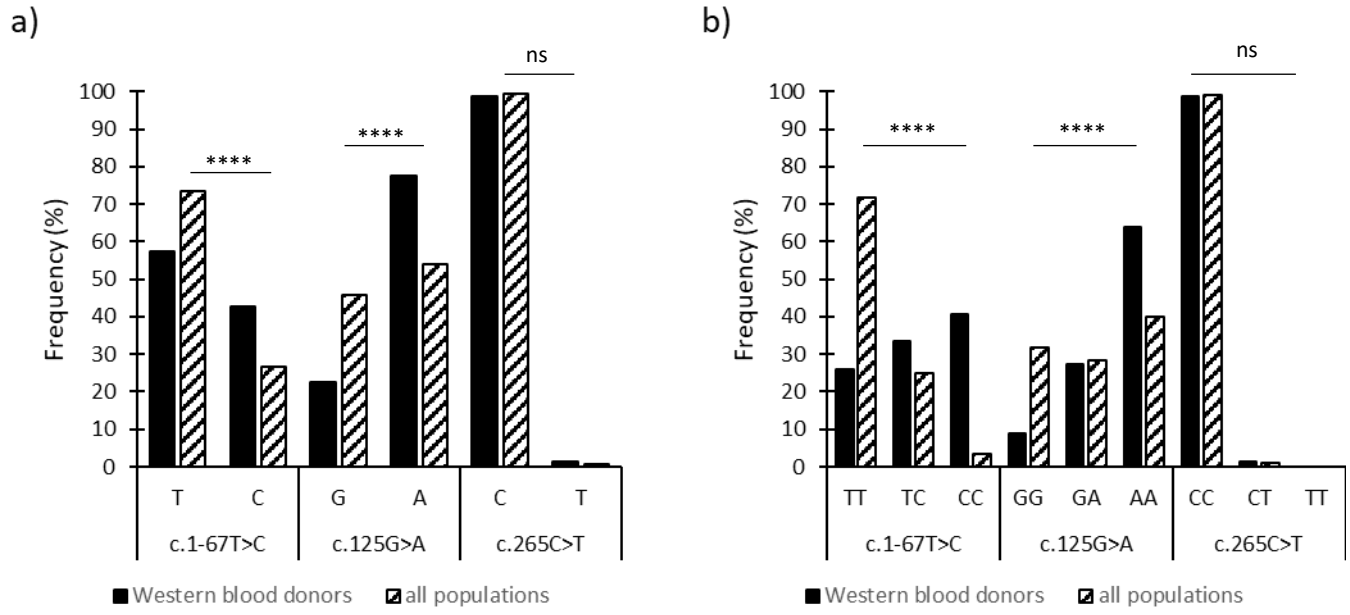

Figure 3 shows the comparisons between the *FY* alleles and genotypes frequencies found in the Western blood donors of Saudi Arabia and the overall alleles and genotypes frequencies of all population on 1000 genomes database. (a) The allele frequencies of the c.1-67T>C, c.125G>A, and the c.265C>T single nucleotide polymorphisms in Western blood donors were significantly different from the overall allele frequncies of all populations (*P – value <0.01*). (b) Similar findings were found between the genotype frequencies in the Western blood donors and the overall genotype frequencies of all populations for all single nucleotide polymorphisms analyzed in this study (*P – value <0.01*). Statistical analyses of the contingency tables were conducted using two-tailed 2X2 or 3X2 chi-squared test to whether there was no difference between the allele and genotype frequency found in this study and the overall alleles and genotype frequencies of all populations found on 1000 genome database. Alternatively, 2X2 fisher exact test was used when the contingency table contains an allele or genotype frequency of less than 5. (*) statistically significant difference.

Table 4 The frequency of c.1-67T>C, c.125G>A, and c.265C>T haplotype combinations in blood donors from Western Saudi Arabia.

Haplotype			Total number of donors N (%)	Predicted genotype	Predicted antigen	Predicted phenotype
c.1-67T>C	c.125G>A	c.265C>T				
TT	AA	CC	14 (4.01)	*FY*02/FY*02*	Fy^b/Fy^b	Fy(a-b+)
TT	GG	CC	28 (8.02)	*FY*01/FY*01*	Fy^a/Fy^a	Fy(a+b-)
TT	GA	CC	45 (12.89)	*FY*01/FY*02*	Fy^a/Fy^b	Fy(a+b+)
CC	AA	CC	135 (38.68)	*FY*02N.01/FY*02N.01*	Fy^{es}/Fy^{es}	Fy(a-b-)
CC	GA	CC	7 (2.01)	*FY*01N.01/ FY*02N.01*	Fy^{es}/Fy^{es}	Fy(a-b-)
TC	AA	CC	71 (20.34)	*FY*02/FY*02N.01*	Fy^b/Fy^{es}	Fy(a-b+)
TC	GA	CC	42 (12.03)	*FY*01/FY*02N.01*	Fy^a/Fy^{es}	Fy(a+b-)
TC	GG	CC	2 (0.57)	*FY*01/ FY*01N.01*	Fy^a/Fy^{es}	Fy(a+b-)
TT	AA	CT*	1 (0.29)	*FY*02/FY*02W.01*	Fy^b/Fy^x	Fy(a-b+)
TT	GA	CT*	1 (0.29)	*FY*01/FY*02W.01*	Fy^a/Fy^x	$Fy(a+b+^w)$
TT	GG	CT*	1 (0.29)	*FY*01/ FY*01W.02*	$Fy^a/Fy^{a(w)*}$	Fy(a+b-)
TC	AA	CT*	2 (0.57)	*FY*02W.01/FY*02N.01*	Fy^x/Fy^{es}	$Fy(a-b+^w)$
Total			349			

All c.265T>C haplotypes in these individuals were associated with the c.298G>A polymorphism

Table 5 Duffy allele frequencies and genotypes in Western Saudi Arabian blood donors in comparison to other population frequencies, obtained from 1000 genomes database.

	Allele	Genotype	Western Saudis No. (%)	All populations No. (%)	*P - Value*
c.1-67T>C	T		401 (57.5)	3674 (73.4)	<0.01
	C		297 (42.5)	1334 (26.6)	
		TT	90 (25.8)	1793 (71.6)	
		TC	117 (33.5)	88 (3.5)	<0.01
		CC	142 (40.7)	623 (24.9)	
c.125G>A	A		541 (77.5)	2707 (54.1)	<0.01
	G		157 (22.5)	2301 (45.9)	
		GG	31 (8.9)	794 (31.7)	<0.01
		GA	95 (27.2)	713 (28.5)	
		AA	223 (63.9)	997 (28.5)	
c.265C>T	C		693 (99.3)	4985 (99.5)	0.36
	T		5 (0.7)	23 (0.5)	
		CC	344 (98.6)	2481 (99.1)	
		CT	5 (1.4)	23 (0.9)	0.38
		TT	0 (0)	0 (0)	

*Either Pearson's chi squared or Fisher Exact test P-value was used to examine differences between allele distributions in the two blood donor groups. No., number of individuals with the allele or genotype.

Duffy haplotype analyses in blood donors from Western Saudi Arabia

We analysed haplotype combinations between the *c.1-67T>C*, *c.125G>A*, and c.256C>T polymorphisms, and their predicted genotypes and phenotypes in samples. Overall, we observed no significant difference in haplotypes between samples from Makkah and Jeddah. The overall frequency of haplotypes in Western blood donors is shown (Table 4 & 5). The homozygous haplotypes, TT GG CC and TT AA CC were found in 8.02 and 4.01% of blood donors, respectively. These haplotypes represented the predicted

genotypes of *FY*01/FY*01* and *FY*02/FY*02*, respectively. The TT GA CC haplotype was found in 45 (12.89%) samples, demonstrating a predicted heterozygous genotype of *FY*01/FY*02*, and a predicted phenotype, Fy(a+b+). The majority of samples (38.68%) had the CC AA CC haplotype, with a predicted homozygous genotype of *FY*02N.01/FY*02N.01*, and a predicted phenotype, Fy(a-b-). The heterozygous genotype, *FY*01N.01/FY*02N.01* was observed in seven donors, and showed an erythroid silent allele association with the *FY*01* allele. The TC AA CC haplotype was found in 20.34% of samples, with a predicted genotype of *FY*02/FY*02N.01*. A considerable proportion of blood donors (12.03%) inherited the TC GA CC haplotype. This haplotype may lead to a predicted genotype of *FY*01/FY*02N.01*, and a predicted phenotype, Fy(a+b-). This was also observed in two samples (0.57%), where the predicted genotype was *FY*01/FY*01N.01*. The association of c.265C>T with *FY*01* or *FY*02* alleles was observed in five samples. The TT AA CT haplotype, with a predicted genotype of *FY*02/FY*02W.01* and a predicted phenotype, Fy(a-b+) was found in a single sample. In addition, a second donor had the predicted *FY*01/FY*02W.01* genotype, and a predicted phenotype of Fy(a+b+^{w}). The predicted genotype *FY*01/FY*01W.02*, where the *FY*X* allele is associated with the FY*01 allele may lead to a predicted phenotype of Fy(a+b-). The remaining two samples demonstrated a predicted genotype of *FY*02W.01/FY*02N.01*, and a predicted phenotype of Fy(a-b+^{w}). No samples had the following haplotype combinations: CC AA CT, CC GA CT and CC GG CT.

4. DISCUSSION

Duffy blood group alleles and genotypes can be used as anthropological markers (Kempińska-Podhorodecka et al., 2012, Howes et al., 2011). In the Saudi Arabian population, blood group genotyping studies have been limited to *ABO* and *RHD* blood group systems. In this study, we aimed to establish a genetic profile of Duffy blood group in Western Saudis by analysing the allele and genotype frequencies of the Duffy blood group in 349 blood donors from Jeddah and Makkah cities. Allele frequencies of blood donors showed a noticeable increase in *FY*B* alleles in Jeddah blood donors, when compared to Makkah blood donors. This was probably due to the presence of an Asian minority in Makkah city, who are expected to have inherited the *FY*A* allele. Nevertheless, we observed no significant differences in genotype frequencies between the two blood donor groups, suggesting our findings may be used as a database for Western Saudis, and for comparisons with other populations. Analysis of the c.125G>A SNP demonstrated a dominance of the AA genotype in the majority of samples, similar to the frequency found in Caucasian populations (Guerra et al., 2009). In addition, *FY*B* is the ancestral allele in Western Saudi Arabians, as previously reported in other populations (McManus et al., 2017).

To a lesser extent, the frequency of the heterozygous GA genotype was similar to the Baoulch population, in South-East Iran (Miri-Moghaddam et al., 2014). The GG genotype was found in only 8.9% of samples, demonstrating marked differences from Asian populations, such as Chinese and Thai, where the GG genotype is predominant (De Silva et al., 2014; Nathalang et al., 2015). However, the majority of samples had the homozygous erythroid silent genotype *c.1-67C*, similar to Africans (Iwamoto et al., 1996), but with a frequency of less than half in our samples. The observation of Duffy polymorphisms disrupted Hardy-Weinberg equilibrium (HWE) in both sample sets (Jeddah and Makkah), except for the *c.256C>T* polymorphism. This suggests that allele and genotype distributions may have changed over time, if the assumptions of the HWE remained disrupted. The analysis of RBC gene divergence in populations could provide insights into the national history, human origins, and migratory patterns, especially in Saudi Arabia where populations vary between geographical locations. Population data from the 1000 genomes database, which included African, American, East Asian, European and South Asian data were compared to the allele and genotype frequencies observed in this study (Table 5, Figure 3a - b). The allele and genotype frequencies of the *c.1-67T>C* polymorphism were different from other populations, including Africans. Similar findings were observed for the *c.125G>A* polymorphism. As for the *c.265C>T* polymorphism, we observed no differences between allele and genotype distributions between our samples and other populations. These data suggest that Western Saudi Arabia may contain a unique genetic pool, in comparison to other populations.

Combined haplotype analyses of Duffy polymorphisms were used to determine the genotypes in Western Saudis. A quarter of Western Saudis inherited *FY*01* (*FY*A*) or *FY*02* (*FY*B*) or both, and thus, expressed Fy antigens are on RBCs whereas the majority of them had the *c.1-67T>C* polymorphism in association with the *FY*02* allele but few were associated with the *FY*01* allele. The genotype *FY*02N.01/FY*02N.01* presented in this work is previously described in Africans, and Brazilians of African ancestry but rarely observed in Caucasians (Tournamille et al., 1995; Castillho et al., 2004). This genotype was less prevalent in Western Saudis in comparison to Africans (Miller et al., 1976; Schmid et al., 2012). Interestingly, the association of the GATA mutation with the *FY*01* allele was also observed in this study as previously reported in Papua New Guinea and Sudan, confirming that the point mutation in the promoter region of the *FY* gene is not restricted to the *FY*02* (*FY*B*) allele in Western Saudis as well as in other populations (Zimmerman et al., 1999; Kempinska-Podhorodecka et al., 2012). However, the *FY*01N.01* was in heterozygous with the *FY*02N.01* allele in our study and its clinical significance will not be different from the *FY*02N.01/FY*02N.01* genotype. Adding to the

heterogeneity of the *FY* alleles and genotypes in Western Saudis, considerable frequencies of the *FY*02N.01* erythroid silencing allele was found in heterozygous with either *FY*02* or *FY*01*. Also, the rare *FY*01/ FY*01N.01* was also reported in our study where the gene- dosage effect was observable in a previous report (Zimmerman et al., 1999). Other *FY* erythroid silencing mutations found in Caucasians and Brazilians such as *c.1-69T>C*, *c.304G>A* and *c.271T>G* were not found in our study (Castillho et al., 2004; Písačka et al., 2015). Genotyping of Western Saudis also demonstrated that few individuals inherited weakened *FY* genotypes. The *c.256T>C* and *c.298G>A* polymorphisms were associated, not only with the *FY*02* allele, but also with the *FY*01* allele which is in agreement with previous reports (Arndt et al., 2015; Lopez et al., 2015). Among these genotypes, the *FY*02W.01/FY*02N.01* is a unique and very rare combination of weak *FY* allele and the erythroid silencing mutation that was found in a Sardinian family (Manfroi et al., 2015). The conformational changes on the expression of duffy glycoproteins on erythrocytes resulting from *FY*02W.01/FY*02N.01* is yet to be elucidated. Individuals with this combination may receive Fy(a-b+) erythrocytes nonetheless, the potential risk of alloimmunisation for patients with this genotype is low as the majority of the Western Saudis have the *FY*Null* genotype. Overall, it is not known how the Western Saudis have mixed *FY* genotypes. It can be assumed that this region is close to Africa, and immigrant pilgrimages and settlements in the main cities of Makkah and Jeddah may have contributed to the genetic admixture. Nationwide genotyping might reveal more insights into the ancestral origins of Saudi Arabians. Obviously, future investigations of *FY* genotypes should encompass single nucleotide polymorphisms that is associated with weak Fy antigens. In fact, whole *FY* gene sequencing would be more informative and may introduce novel SNPs in different populations.

Duffy blood group antigens and alloantibodies have clinical implications. Duffy alloantibodies are implicated in both immediate and delayed transfusion reactions (Daniels, 2013). In our study, approximately 40% of samples were homozygous for the GATA mutation, suggesting resistance to *P.vivax* invasion, as previously reported (Hedrick et al., 2011). However, those individuals with rarely develop anti-Duffy alloanitbodies as Duffy antigens are also expressed in non-erythroid cells (kosinki). In contrast, individuals expressing Duffy antigens on erythrocytes may receive incompatible blood, and develop anti-Duffy alloantibodies. In our research, the majority of blood donors expressed Duffy antigens, and few had a weak Duffy genotype. This observation is important for Makkah, as over two million pilgrims from several ethnic origins visit the city annually. For this event, the local authorities must collect and reserve blood from blood banks, including those in Jeddah and Makkah cities, for emergencies. Data on pilgrim alloimmunisation and transfusion reactions are unavailable. Hypothetically, alloimmunisation to Duffy antigens and other blood groups may occur in these pilgrims. On a smaller scale, a recent alloimmunisation study in sickle cell and thalassaemia patients demonstrated the presence of anti-Duffy alloantibodies, and alloantibodies for other blood groups (Hindawi et al., 2020). Therefore, implementing molecular genotyping, combined with serological techniques in providing compatible blood units, may reduce alloimmunisation risks and resolve complex phenotyping cases in Western Saudis and pilgrims as previously proposed (Jungbauer, 2011).

5. CONCLUSION

To our knowledge, Duffy blood group alleles and genotypes have not been previously investigated in Saudi Arabia. The molecular genetic background of Western Saudis is, to some extent, similar to Africans, probably due to geographical proximity and immigration. This complex situation has resulted in a unique Duffy allele and genotype distribution from other populations, and may have clinical implications including the risk of alloimmunisation for pilgrims visiting Makkah. Thus, introducing molecular genotyping as a standard practice for blood banks in Saudi Arabia may reduce alloimmunisation risks.

Acknowledgment

We would like to thank Oxbridge proofreading (https://oxbridgeproofreading.co.uk) for language editing.

Author Contributions

WMB conceived and designed the study. AA, MS, RA, RF, NH, MA, AM, RE, and TS conducted research, provided research materials, and collected and organized data. WMB and AK analyzed and interpreted data. All authors wrote initial and final draft of the article. All authors have critically reviewed and approved the final draft of the manuscript.

Funding

This study has not received any external funding.

Competing interests

The authors declare that they have no competing interests.

Informed consent

Written and Oral informed consent was obtained from all individual participants included in the study. Additional informed consent was obtained from all individual participants for whom identifying information is included in this manuscript.

Ethical Approval

The research was conducted according to the declaration of Helsinki. The Institutional Review Board (IRB) at the King Fahad City (Riyadh, Saudi Arabia) approved sample collection and analysis (IRB log number: 16-459E).

Data and materials availability

All data associated with this study are present in the paper.

REFERENCES AND NOTES

1. Alabdulmonem W, Shariq A, Alqossayir F, AbaAlkhail FM, Al-Musallam AY, Alzaaqi FO, Aloqla AA. Sero-prevalence ABO and Rh blood groups and their associated Transfusion-Transmissible Infections among Blood Donors in the Central Region of Saudi Arabia. J Infect Public Health 2020; 13: 299–305.
2. Arndt PA, Horn T, Keller JA, Heri SM, Keller MA. First example of an FY*01 allele associated with weakened expression of Fya on red blood cells. Immunohematology 2015; 31: 103–107.
3. Bashwari LA, Al-Mulhim AA, Ahmad MS, Ahmed MA. Frequency of ABO blood groups in the Eastern region of Saudi Arabia. Saudi Med J 2001; 22: 1008–1012.
4. British Committee for Standards in Haematology, Milkins C, Berryman J, Cantwell C, Elliott C, Haggas R, et al. Guidelines for pre-transfusion compatibility procedures in blood transfusion laboratories. British Committee for Standards in Haematology. Transfus Med 2013; 23: 3–35.
5. Castilho L, Rios M, Pellegrino Jr J, Saad STO, Costa FF, Reid ME. A novel FY allele in Brazilians. Vox Sanguinis 2004; 87: 190–195.
6. Chaudhuri A, Zbrzezna V, Johnson C, Nichols M, Rubinstein P, Marsh WL, Pogo AO. Purification and characterization of an erythrocyte membrane protein complex carrying Duffy blood group antigenicity. Possible receptor for Plasmodium vivax and Plasmodium knowlesi malaria parasite. J Biol Chem 1989; 264: 13770–13774.
7. Chitnis CE, Chaudhuri A, Horuk R, Pogo AO, Miller LH. The domain on the Duffy blood group antigen for binding Plasmodium vivax and P. knowlesi malarial parasites to erythrocytes. J Exp Med 1996; 184: 1531–1536.
8. Cutbush M, Mollison PL, Parkin DM. A New Human Blood Group. Nature 1950b; 165: 188–189.
9. Cutbush M, Mollison PL. The Duffy blood group system. Heredity (Edinb) 1950a; 4: 383–389.
10. Daniels G. Duffy Blood Group System. In: Human Blood Groups [Internet]. 2013. John Wiley & Sons, Ltd; [Cited 2020 Aug 9]. p. 306–24. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118493595.ch8
11. Daniels G. The molecular genetics of blood group polymorphism. Transpl Immunol 2005; 14: 143–153.
12. De Silva JR, Lau YL, Fong MY. Genotyping of the Duffy blood group among Plasmodium knowlesi-infected patients in Malaysia. PLoS One 2014; 9: e108951.
13. Guerra CA, Howes RE, Patil AP, Gething,PW, Van Boeckel TP, Temperley WH, et al. The international limits and population at risk of Plasmodium vivax transmission in 2009. PLoS Negl Trop Dis 2010; 4: e774.
14. Hadley TJ, Klotz FW, Miller LH. Invasion of Erythrocytes by Malaria Parasites: A Cellular and Molecular Overview. Annu Rev Microbiol 1986; 40: 451–477.
15. Hedrick PW. Population genetics of malaria resistance in humans. Heredity (Edinb) 2011; 107: 283–304.
16. Hindawi S, Badawi M, Elfayoumi R, Elgemmezi T, Al Hassani A, Raml M, Alamoudi S, Gholam K. The value of transfusion of phenotyped blood units for thalassemia and sickle cell anemia patients at an academic center. Transfusion 60 Suppl 2020; 1: S15–S21.
17. Höher G, Fiegenbaum M, Almeida S. Molecular basis of the Duffy blood group system. Blood Transfus 2018; 16: 93–100.
18. Howes RE, Patil AP, Piel FB, Nyangiri OA, Kabaria CW, Gething PW, et al. The global distribution of the Duffy blood group. Nat Commun 2011; 2: 266.
19. Hughes LH, Rossi KQ, Krugh DW, O'Shaughnessy RW. Management of pregnancies complicated by anti-Fya alloimmunization. Transfusion 2007; 47: 1858–1861.
20. Iwamoto S, Li J, Sugimoto N, Okuda H, Kajii E. Characterization of the Duffy gene promoter: evidence for tissue-specific abolishment of expression in Fy(a-b-) of black

individuals. Biochem Biophys Res Commun 1996; 222: 852–859.
21. Jungbauer C. Blood group molecular genotyping. ISBT Science Series 2011; 6: 399–403.
22. Kempińska-Podhorodecka A, Knap O, Drozd A, Kaczmarczyk M, Parafiniuk M, Parczewski M, Ciechanowicz A. Analysis for genotyping Duffy blood group in inhabitants of Sudan, the fourth cataract of the Nile. Malar J 2012; 11: 115.
23. Kim HH, Park TS, Oh SH, Chang CL, Lee EY, Son HC. Delayed hemolytic transfusion reaction due to anti-Fyb caused by a primary immune response: a case study and a review of the literature. Immunohematology 2004; 20: 184–186.
24. Le Pennec PY, Rouger P, Klein MT, Robert N, Salmon C. Study of anti-Fy^a in five black Fy(a-b-) patients. Vox Sang 1987; 52: 246–249.
25. Lee CS, Ying KL, Bowen P. Position of the Duffy locus on chromosome 1 in relation to breakpoints for structural rearrangements. Am J Hum Genet 1974; 26: 93–102.
26. Lopez GH, Condon JA, Wilson B, Martin JR, Liew Y-W, Flower RL, Hyland CA. A novel FY*A allele with the 265T and 298A SNPs formerly associated exclusively with the FY*B allele and weak Fy(b) antigen expression: implication for genotyping interpretative algorithms. Vox Sang 2015; 108: 52–57.
27. Luo H, Chaudhuri A, Zbrezezna V, He Y, Pogo AO. Deletion of Murine Duffy Gene (Dfy) Reveals that the Duffy Receptor Is Functionally Redundant. Mol Cell Biol 2003; 23: 3030–3030.
28. Manfroi S, Scarcello A, Pagliaro P. Molecular identification of rare FY*Null and FY*X alleles in Caucasian thalassemic family from Sardinia. Transfus Apher Sci 2015; 53: 225–227.
29. McManus KF, Taravella AM, Henn BM, Bustamante CD, Sikora M, Cornejo OE. Population genetic analysis of the DARC locus (Duffy) reveals adaptation from standing variation associated with malaria resistance in humans. PLoS Genet 2017; 13: e1006560.
30. Meny GM. The Duffy blood group system: a review. Immunohematology 2010; 26: 51–56.
31. Miller LH, Mason SJ, Clyde DF, McGinniss MH. The resistance factor to Plasmodium vivax in blacks. The Duffy-blood-group genotype, FyFy. N Engl J Med 1976; 295: 302–304.
32. Miri-Moghaddam E, Bameri Z, Mohamadi M. Duffy blood group genotypes among malaria Plasmodium vivax patients of Baoulch population in Southeastern Iran. Asian Pac J Trop Med 2014; 7: 206–207.
33. Mohandas N. Blood group antigens in health and disease. Curr Opin Hematol 2005; 12: 135–140.
34. Nathalang O, Intharanut K, Siriphanthong K, Nathalang S, Kupatawintu P. Duffy blood group genotyping in Thai blood donors. Ann Lab Med 2015; 35: 618–623.
35. Olsson ML, Smythe JS, Hansson C, Poole J, Mallinson G, Jones J, Avent ND, Daniels G. The Fy(x) phenotype is associated with a missense mutation in the Fy(b) allele predicting Arg89Cys in the Duffy glycoprotein. Br J Haematol 1998; 103: 1184–1191.
36. Olteanu H, Gerber D, Partridge K, Sarode R. Acute hemolytic transfusion reaction secondary to anti-Fy3. Immunohematology 2005; 21: 48–52.
37. Owaidah AY, Naffaa NM, Alumran A, Alzahrani F. Phenotype Frequencies of Major Blood Group Systems (Rh, Kell, Kidd, Duffy, MNS, P, Lewis, and Lutheran) Among Blood Donors in the Eastern Region of Saudi Arabia. J Blood Med 2020; 11: 59–65.
38. Pisacka M, Vytiskova J, Latinakova A, Koristka M, Strasikova M. Molecular background of the Fy(ab) phenotype in Gypsy population living in the Czech and Slovak Republics. Transfusion 2001; 41: 15S.
39. Pogo AO, Chaudhuri A. The duffy protein: A malarial and chemokine receptor. Seminars in Hematology, Blood Group Antigens: Molecular Biology, Functions, and Clinical Applications 2000; 37: 122–129.
40. Reesink HW, Engelfriet CP, Schennach H, Gassner C, Wendel S, Fontão-Wendel R, et al. Donors with a rare pheno (geno) type. Vox Sang 2008; 95: 236–253.
41. Reid ME, Mohandas N. Red blood cell blood group antigens: structure and function. Semin Hematol 2004; 41: 93–117.
42. Rios M, Chaudhuri A, Mallinson G, Sausais L, Gomensoro-Garcia AE, Hannon J, Rosenberger S, Poole J, Burgess G, Pogo O, Reid M. New genotypes in Fy(a-b-) individuals: nonsense mutations (Trp to stop) in the coding sequence of either FY A or FY B. Br J Haematol 2000; 108: 448–454.
43. Rot A, von Andrian UH. Chemokines in Innate and Adaptive Host Defense: Basic Chemokinese Grammar for Immune Cells. Annu Rev Immunol 2004; 22: 891–928.
44. Schmid P, Ravenell KR, Sheldon SL, Flegel WA. DARC alleles and Duffy phenotypes in African Americans. Transfusion 2012; 52: 1260–1267.
45. Talano J-A, Hillery CA, Gottschall JL, Baylerian DM, Scott JP. Delayed Hemolytic Transfusion Reaction/ Hyperhemolysis Syndrome in Children with Sickle Cell Disease. Pediatrics 2003; 111: e661–e665.
46. Zimmerman PA, Woolley I, Masinde GL, Miller SM, McNamara DT, Hazlett F, et al. Emergence of FY*A(null) in a Plasmodium vivax-endemic region of Papua New Guinea. Proc Natl Acad Sci USA 1999; 96: 13973–13977.